Reading Text — Stage 01: Problem Framing & Scoping

Why Framing Comes First

Financial engineering starts from real-world questions. Before data and models, define **who** will use your work, **what decision** it informs, and **what form** of answer helps them. This avoids "solution in search of a problem."

Stakeholder-Centered Analytics

- **Stakeholder vs. User:** The stakeholder owns the decision; the user may operate the tool. They can be the same person—but often are not.
- **Decision Window:** Identify when the decision is made (daily, weekly, monthly). This sets latency and refresh requirements.
- **Useful Answer Types:**
    - **Descriptive:** What happened? Dashboards, summaries, anomaly flags.
    - **Predictive:** What will happen? Forecasts, rankings, risk bands.
    - **Causal:** What action changes outcomes? Policy/treatment effects, A/B tests.

Assumptions and Constraints

State assumptions explicitly (e.g., stationarity, liquidity, stable relationships). Add constraints like capacity, compliance, runtime, and data latency. Good assumptions are **testable**; plan how to validate them later.

Known Unknowns (Risks)

Surface uncertainty early: missing data, survivorship bias, model instability, regime shifts. Design monitoring hooks to revisit assumptions as evidence accumulates.

From Problem to Repo

A minimal, collaborative repo on day one keeps work organized and usable:

```
/data/        # raw/processed data (avoid pushing sensitive data)
/src/         # reusable functions and scripts
/notebooks/   # EDA, experiments (numbered)
/docs/        # memos, personas, design notes
README.md     # scoping paragraph + mapping
```

Write a README first. It is your contract with future collaborators.

Mapping Goals → Lifecycle → Deliverables

Create a simple mapping that ties each goal to a stage and an artifact. Example:

- **Goal:** Inform weekly ETF rebalance → **Stage:** Problem Framing & Scoping → **Deliverable:** Scoping paragraph + persona + repo skeleton.

Common Misconceptions

- **"We'll fix it in modeling."** No: a fuzzy question yields fragile models.
- **"Data will be there."** Often not; plan for gaps and negotiation with data owners.
- **"Users = stakeholders."** Sometimes; verify roles and incentives.

Mini Case Walkthrough (ETF Strategy)

- Question: "Will this ETF strategy still work next year?"
- Decision owner: PM; users: PM + analyst.
- Useful answer: predictive intervals + plain-language scenarios.
- Assumptions: stationarity, liquidity, realistic capacity.
- Risks: regime change; cost drift; benchmark mismatch.

Checklist for Your Paragraph

- ☐ Specific problem and why it matters
- ☐ Named stakeholder and (if different) user
- ☐ Useful answer type and decision trigger
- ☐ Assumptions & constraints
- ☐ Known unknowns / risks
- ☐ Lifecycle mapping and minimal repo plan